

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comagen Ltd.

OW protein - protein search, using sw model

Run on: December 15, 2003, 16:51:32 / Search time 37 Seconds  
(without alignments)  
2397.681 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457  
Sequence: 1 MPEDEPENTQPLIGPPGSA.....CVPEKTKGTEIGITMPEGR 477

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Maximum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /cgn2\_6/prodata/1/pubppa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/prodata/1/pubppa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/prodata/1/pubppa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/prodata/1/pubppa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/prodata/1/pubppa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/prodata/1/pubppa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/prodata/1/pubppa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/prodata/1/pubppa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/prodata/1/pubppa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/prodata/1/pubppa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/prodata/1/pubppa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/prodata/1/pubppa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/prodata/1/pubppa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/prodata/1/pubppa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/prodata/1/pubppa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/prodata/1/pubppa/US60\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep:\*
- 18: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2457	100.0	477	US-10-168-651-1	Sequence 1, Appli
2	948	38.6	507	US-10-157-031-359	Sequence 359, App
3	723	29.4	501	US-10-051-909-38	Sequence 38, Appl
4	686.3	26.8	533	US-10-051-909-38	Sequence 36, Appl
5	615	25.0	488	US-10-094-059-4	Sequence 4, Appl1
6	615	25.0	488	US-10-170-528-5	Sequence 5, Appl1
7	615	25.0	488	US-10-162-012-46	Sequence 46, Appl1
8	615	25.0	488	US-10-062-9608-4	Sequence 4, Appl1
9	615	25.0	487	US-10-144-624-4	Sequence 27, Appl1
10	613.5	25.0	487	US-10-391-399-45	Sequence 45, Appl1
11	613.5	25.0	487	US-10-095-139-14	Sequence 14, Appl1
12	613.5	25.0	487	US-10-156-239-17	Sequence 27, Appl1
13	613.5	25.0	487	US-10-199-485-27	Sequence 27, Appl1
14	613.5	25.0	472	US-10-154-419-79	Sequence 79, Appl1
15	603	24.5	472	US-10-146-733-74	Sequence 74, Appl1

Result No.	Score	Query Match	Length	DB ID	Description
16	603	24.5	472	US-10-146-733-74	Sequence 74, Appl1
17	603	24.5	472	US-10-024-623-29	Sequence 29, Appl1
18	598	24.3	486	US-09-860-232A-7	Sequence 7, Appl1
19	591.5	24.1	464	US-10-154-419-78	Sequence 78, Appl1
20	591.5	24.1	464	US-10-116-733-73	Sequence 73, Appl1
21	591.5	24.1	464	US-10-024-623-28	Sequence 28, Appl1
22	588.5	24.0	535	US-09-795-693-20	Sequence 20, Appl1
23	588.5	24.0	535	US-10-156-239-20	Sequence 20, Appl1
24	588.5	24.0	535	US-10-159-485-20	Sequence 20, Appl1
25	575.5	23.4	494	US-09-981-947A-5	Sequence 5, Appl1
26	570.5	23.2	496	US-10-341-434-89	Sequence 89, Appl1
27	568	23.1	493	US-09-981-947A-10	Sequence 10, Appl1
28	548.5	22.3	571	US-10-051-909-34	Sequence 34, Appl1
29	543	22.1	523	US-10-099-332-122	Sequence 122, Appl
30	542.5	22.1	522	US-10-099-332-125	Sequence 125, App
31	541.5	22.0	522	US-10-099-332-124	Sequence 124, App
32	540	22.0	523	US-10-099-332-123	Sequence 123, App
33	536	21.8	509	US-09-981-947A-6	Sequence 6, Appl1
34	535	21.7	519	US-10-259-185-248	Sequence 248, App
35	529	21.5	492	US-09-205-658-208	Sequence 208, App
36	528	21.5	492	US-09-963-693-208	Sequence 208, App
37	522	21.2	617	US-10-154-419-52	Sequence 52, Appl
38	522	21.2	617	US-10-146-733-47	Sequence 47, Appl1
39	522	21.2	617	US-10-024-623-2	Sequence 2, Appl1
40	521.5	21.2	471	US-09-778-927A-17	Sequence 17, Appl1
41	521.5	21.2	501	US-09-778-927A-48	Sequence 48, Appl1
42	519	21.1	491	US-09-728-626-5703	Sequence 3703, Ap
43	516.5	21.0	524	US-10-328-198-3	Sequence 3, Appl1
44	513	20.9	524	US-10-099-332-34	Sequence 34, Appl1
45	513	20.9	524	US-10-099-332-121	Sequence 121, App

ALIGNMENTS

RESULT 1  
US-10-168-651-1  
US-10-168-651-1, Application US/10168651  
Publication No. US20030171275A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: BAUGHAN, Mariah R.  
APPLICANT: BURFORD, Neil  
APPLICANT: AU-YOUNG, Janice  
APPLICANT: LU, Dyoung Ahna M.  
APPLICANT: YANG, Junming  
APPLICANT: RADDY, Koopa  
APPLICANT: LAU, Preeti  
APPLICANT: HILLMAN, Jennifer L.  
APPLICANT: AZIMZAI, Yalda  
APPLICANT: YUE, Henry  
APPLICANT: NGUYEN, Damien B.  
APPLICANT: VAO, Monique G.  
APPLICANT: GANDHI, Ameesh R.  
APPLICANT: TANG, Y. Tom  
APPLICANT: KHAN, Farrah A.  
TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS  
FILE REFERENCE: PI-0005 PCT  
CURRENT APPLICATION NUMBER: US/10/168,651  
PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,322; 60/178,572; 60/179,71  
60/181,625  
PRICOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;  
2000-02-10  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PERL Program  
SEQ ID NO 1  
LENGTH: 477  
TITER: FRT  
ORGANISM: Homo sapiens  
FEATURES:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030171275A1 1416107CD1

US-10-158-651-1

Query Match 100.0%; Score 2457; DB 12; Length 477; Best Local Similarity 100.0%; Pred. No. 1.4e-208; Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MTPEDPEETOPLLGPPGSGNPRGRVPLAFAAALGPPSPFALGYSPPAIESIORAAP 60
1 MTPEDPEETOPLLGPPGSGNPRGRVPLAFAAALGPPSPFALGYSPPAIESIORAAP 60
61 APRLDAAASMPGAVVTLGAAAGVTLGKMLVDRAGKLSILLCSVPFAGAVVTAADV 120
61 APRLDAAASMPGAVVTLGAAAGVTLGKMLVDRAGKLSILLCSVPFAGAVVTAADV 120
61 APRLDAAASMPGAVVTLGAAAGVTLGKMLVDRAGKLSILLCSVPFAGAVVTAADV 120
121 MLLGSRLLTGLACGVAASVAVPYISEIAYPAVRGILGSCVQDMVAVGILLAYLAGWLE 180
121 MLLGSRLLTGLACGVAASVAVPYISEIAYPAVRGILGSCVQDMVAVGILLAYLAGWLE 180
121 MLLGSRLLTGLACGVAASVAVPYISEIAYPAVRGILGSCVQDMVAVGILLAYLAGWLE 180
181 WRMLAVGCVPPSMLMLMCMETPRILLTORRQEMAAALRPLMGSBOGMEPPIGAE 240
181 WRMLAVGCVPPSMLMLMCMETPRILLTORRQEMAAALRPLMGSBOGMEPPIGAE 240
181 WRMLAVGCVPPSMLMLMCMETPRILLTORRQEMAAALRPLMGSBOGMEPPIGAE 240
241 QSFHALIRPGIYKPEIIGVSIAMAFQOLSGVAVVAFYAEIIFEBKPFDSLSAVVVG 300
241 QSFHALIRPGIYKPEIIGVSIAMAFQOLSGVAVVAFYAEIIFEBKPFDSLSAVVVG 300
241 QSFHALIRPGIYKPEIIGVSIAMAFQOLSGVAVVAFYAEIIFEBKPFDSLSAVVVG 300
301 IQVLFVAALIMDRARERLLVLSGVVWVSTASAFAYEKLTOGGPPGNSPVALSAPVS 360
301 IQVLFVAALIMDRARERLLVLSGVVWVSTASAFAYEKLTOGGPPGNSPVALSAPVS 360
301 IQVLFVAALIMDRARERLLVLSGVVWVSTASAFAYEKLTOGGPPGNSPVALSAPVS 360
361 AQPVDASVGLAMLVAGSMCLFJAGFAVGMPIWMLMSEIFPLHWKAVATGICVLTNWM 420
361 AQPVDASVGLAMLVAGSMCLFJAGFAVGMPIWMLMSEIFPLHWKAVATGICVLTNWM 420
421 AFVTKPESSIMELRBYGAFMLASAFCSVLTFLKCPKPKKTLKQITTHPEGR 477
421 AFVTKPESSIMELRBYGAFMLASAFCSVLTFLKCPKPKKTLKQITTHPEGR 477
421 AFVTKPESSIMELRBYGAFMLASAFCSVLTFLKCPKPKKTLKQITTHPEGR 477

RESULT 2

US-10-157-031-359
Sequence 359, Application US/10157031
Publication No. US20030108890A1
GENERAL INFORMATION:
APPLICANT: Baranova, A. V.
APPLICANT: Yanovsky, N. K.
APPLICANT: Kozlov, A. P.
APPLICANT: Lobashev, A. V.
APPLICANT: Krukovskaya, L. L.
TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
FILE REFERENCE: 2760-103
CURRENT APPLICATION NUMBER: US/10/157, 031
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 415
SOFTWARE: Patent in version 3.1
SEQ ID NO 359
LENGTH: 507
TYPE: PRF
ORGANISM: Homo sapiens
US-10-157-031-359

Query Match 38.6%; Score 948; DB 15; Length 507; Best Local Similarity 42.7%; Pred. No. 3e-75; Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

10 QPLG-----PPGGSAPRGR-----RYGLAFAAALGPTLGRFALGYSPPA 50
3 EPLUGAEGPVDTEPEKPEPPGGRARAVGTLQNKRVFLATFAAVLGNRFGLAVYTSV 62
51 IPSLORAAAPRRLDDAASVFGAVVTLGAAAGVTLGKMLVDRAGKLSILLCSVPFAG 110
63 IVALERSIDPDLHLTKSQASVFGSVFTLGAAGGLSMILMINDLGRKLSIMFSAVPSAAG 122

NOY

111 FAVITAAQDVWMLLGRLLTGLACGVAASVAVPYISEIAYPAVRGILGSCVQDMVAVGILL 170
123 YALMAAGHGMWMLLGRLLTGLACGVAASVAVPYISEIAYPAVRGILGSCVQDMVAVGILL 182
171 LAYLAGWLEWRMLAVGCVPPSMLMLMCMETPRILLTORRQEMAAALRPLMGSBOGMEPP 230
183 SLVAGLMLPWRMLAVGCVPPSMLMLMCMETPRILLTORRQEMAAALRPLMGSBOGMEPP 242
221 G--MEDPFG--AEQSFHL--ALLRQGIYKPEIIGVSIAMAFQOLSGVAVVAFYAEIIE 203
243 DVHWBFEQIQNNVARKQSSRVSMABARPHVCRIYALMLRLLQQLTGLITPILVYQSF 302
284 EBAKF---KSSLASVYVYIQLPFAVAALIMDRAGRERLLVLSGVVWVSTASAFAYE 339
303 DSTAVLPPKDD--AAIVGVRMLSVYLAALTMVDAGRRKLLFVSNALMFAANLTLGXY 359
340 FRLTGGPNSHVAISAPVS-----AQPVDASVGLAMLVAGSMCLFJAGFAVGMPI 392
360 IHF---GPRPLSPNSTGLSESSWGDLAQPLAPAGVTLVPLATMLFMGYAVWGSII 416
393 PWLMSSEIFPLHWKAVATGICVLTNWMALVTKKESLMEVLRPFGAFMLASAFICFSV 452
417 TWLMSSEVLPFRAGVASGLCVLMSWTLAVLTKSFLUPVSTFGIQVPPFFAALCVSL 476
453 LFTLFCVPEYKTKLEQITTHPE--GR 477
477 VFTGCCVPEYKTKLEQITTHPE--GR 502

RESULT 3

US-10-051-909-38
Sequence 38, Application US/10051909
Publication No. US20020199217A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Helenczaris, Tim
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: B01163 US CIP
CURRENT APPLICATION NUMBER: US/10/051, 909
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083, 044
PRIOR FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Microsoft Office 97
SEQ ID NO 38
LENGTH: 501
TYPE: PRF
ORGANISM: Oryza sativa
US-10-051-909-38

Query Match 29.4%; Score 723; DB 14; Length 501; Best Local Similarity 35.6%; Pred. No. 2.2e-55; Matches 173; Conservative 89; Mismatches 172; Indels 52; Gaps 8;

17 GGSAPRGRVFLAFA-----AALGRLSFGALGYSPAPRSLQRAAPPAP 62
39 GGGGSGSRLSSASLSDSSVSAVLCVLAALGPIGFGTGGSSST-----QDAILISDL 94
63 RUDDAASVFGAVVTLGAAAGVTLGKMLVDRAGKLSILLCSVPFAGAVVTAADV 122
95 GILSFSFSGSLSVNGVMVGAISGGQIAEYIGRKGSLMIAAIPNIIIGMIALSFAADSSR 154
123 LUGGRLLTGLACGVAASVAVPYISEIAYPAVRGILGSCVQDMVAVGILLAYLAGWLE 182
155 LHWGRLLBFGVGVSVVPPYIAFIARQTKRGLSSVNDLSVITGILLAYLGMVPPR 214
183 WNAVIGCVPSMLMLMCMETPRILLTORRQEMAAALRPLMGSBOGMEPPIGAE-- 240
215 IISVIGILPCLSIILIGLFFIIPSPFWLAKMKMKMDFFESSIQVLRGFE-----TDIIVEVW 269